Technical Report 712

Research Integration: An Essential for Military Psychological Research

Laurel W. Oliver

Basic Research

U. S. Army

Research Institute for the Behavioral and Social Sciences

June 1986

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

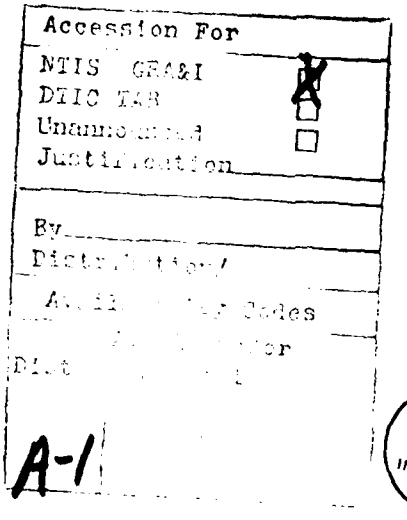
A Field Operating Agency under the Jurisdiction of the
Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Technical review by

George W. Lawton
Joel M. Savell



NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POT, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ARI Technical Report 712	2. GOVT ACCESSION NO.: AD-A177807
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): RESEARCH INTEGRATION: AN ESSENTIAL FOR MILITARY PSYCHOLOGICAL RESEARCH	5. TYPE OF REPORT & PERIOD COVERED: June 1983-June 1985
	6. PERFORMING ORG. REPORT NUMBER: --
7. AUTHOR(s): Laurel W. Oliver	8. CONTRACT OR GRANT NUMBER(s): --
9. PERFORMING ORGANIZATION NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2T161101A91B 4512 100
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	12. REPORT DATE: June 1986
	13. NUMBER OF PAGES: 26
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): --	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

An earlier version of this paper was presented at the Ninth Biennial Psychology in the Department of Defense Symposium, Colorado Springs, April 1984.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Research integration
Research synthesis
Meta-analysis
Military psychological research

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This article encourages military psychologists to learn about and apply nontraditional (quantitative) approaches to research integration in topical areas of interest to military psychology. Various research integration approaches (literary, vote-counting, combining significance levels, and meta-analytic) are summarized. Also discussed are the advantages of meta-analytic approaches to research integration as well as some of the problems posed by the use of such quantitative procedures. Implications for military (Continued)

ARI Technical Report 712

20. (Continued)

psychology include more complete reporting of research results, integration of research in areas of interest to military psychologists, and identification of research gaps. Cross-service collaborative efforts are urged to accomplish the research integration step and to plan programmatic research to fill the gaps in our cumulative knowledge.

Technical Report 712

Research Integration: An Essential for Military Psychological Research

Laurel W. Oliver

Basic Research
Milton S. Katz, Director

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

June 1986

Army Project Number
2T161101A91B

Basic Research

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

For decades, psychologists have conducted research in military settings and with military populations. Yet relatively little integration of psychological research has been accomplished in areas of interest to military psychologists. This report acquaints military psychologists with some quantitative approaches to research integration and encourages them to apply these innovative techniques. The author proposes a cross-service, collaborative effort to integrate research of special interest to the military. Such an effort is seen to have great potential payoff for future military psychological research.

EDGAR M. JOHNSON
Technical Director

RESEARCH INTEGRATION: AN ESSENTIAL FOR MILITARY PSYCHOLOGICAL RESEARCH

EXECUTIVE SUMMARY

Requirement:

Relatively little quantitative integration of research has been accomplished in topical areas of interest to military psychologists. Yet there is a great need for integrating the results of military psychological research in order to summarize these findings and make them of maximum use to the military community.

Procedure:

The author summarizes four general techniques for research integration and discusses the advantages of meta-analytic approaches as well as problems encountered in the use of such procedures.

Findings:

Although meta-analytic approaches have great advantages, especially for large bodies of research literature or research which contains conflicting findings, they are not without their problems. Some of these problems are: obtaining unpublished studies, coding difficulties, the nonindependence of effect sizes from the same study, how to handle outliers, and what to do in cases of missing data. The implications of these findings for military psychology include more complete reporting of research results and quantitative integration of research in areas of interest to military psychologists. Such research integration would summarize what has previously been learned and identify areas where more research is needed. The author proposes collaborative efforts by military psychologists across the services to accomplish the needed research integration.

Utilization:

This report provides an introduction to the meta-analytic approach to research integration, and encourages military psychologists to apply this approach in appropriate research areas.

RESEARCH INTEGRATION: AN ESSENTIAL FOR MILITARY PSYCHOLOGICAL RESEARCH

CONTENTS

RESEARCH INTEGRATION: AN ESSENTIAL FOR MILITARY PSYCHOLOGICAL RESEARCH

INTRODUCTION

The number of research studies has rapidly increased in many disciplines, including psychology. Yet this proliferation of research seems not to have advanced the state of science to the extent that one might have expected. In considering the problem of so much research and the relatively few conclusions that could be drawn from it, Frank Schmidt concluded that "the most important problem in psychology and the social sciences today is the failure to produce cumulative knowledge" (Schmidt, 1980).

Over the past few years, researchers have responded to Schmidt's challenge and become increasingly concerned about the problem of producing "cumulative knowledge." More innovative approaches for integrating research have been proposed and, more importantly, used. These innovative techniques enable the reviewer to combine research findings in a systematic, quantitative manner. While quantitative techniques for integrating research results are not new (e.g., Mosteller & Bush, 1954), until recently they have been infrequently used.

Within the psychological arena lies a sizable body of research of interest to military psychologists. For decades, a wide variety of psychological research has been conducted in military settings and with military populations (Oliver, 1984). But to date relatively little quantitative research integration has been accomplished in areas of interest to military psychologists. A notable exception to this generalization is the integration of research on Fiedler's contingency model of leader effectiveness (Peters, Hartke, & Pohlmann, 1985; Strube & Garcia, 1981).

The purpose of this paper is to acquaint military psychologists with some quantitative (with emphasis on meta-analytic) approaches to research integration and to encourage them to apply such techniques in appropriate research areas. Accordingly, this paper will summarize various approaches to research integration, discuss some advantages and problems associated with using meta-analytic approaches to research integration, and suggest some implications for military psychological research.

APPROACHES TO RESEARCH INTEGRATION

Most research reviews are what Cooper (1984) has called "integrative" research reviews,[1] whose purpose is to present overall summary statements of the findings of a body of research in the topical area. Several approaches that can be used to integrate research results are summarized below.

---

[1]The other types of research reviews are: (1) theoretical reviews, which compare a set of theories, and (2) methodological reviews, which critique research methods (Cooper, 1984).

Literary Approach

The literary or narrative approach has been the traditional procedure for integrating research findings. The reviewer reads the studies on a given topic and attempts to derive generalizations that summarize the results of those studies and thus make the findings useful to the reader. This can be an extraordinarily difficult task for the reviewer. For a sizable body of research, especially one with conflicting findings, the reviewer may experience a severe cognitive overload. Because of the subjective nature of the process, reviewers working with essentially the same set of studies may come to different conclusions (Smith, Glass, & Miller, 1980). Additional bias may be introduced if the reviewer has excluded certain studies which he or she considers inadequate--usually on methodological grounds. Narrative reviews typically end with a call for more research to resolve the confusing findings.

Vote-Counting Approach

In the vote-counting or box-score approach, a directional hypothesis is assumed. The reviewer classifies the findings of each study into positive significant, negative significant, and nonsignificant categories and tallies the results. (Reviewers typically do not indicate the direction of the nonsignificant results.) Vote-counting is more systematic and perhaps more revealing than the narrative review, but this approach also has its drawbacks. The ambiguity of the results may make it difficult for the reviewer to draw firm conclusions about the research being integrated.

There are other disadvantages to the vote-counting approach. Because it is based only on statistical significance, vote counting does not take into account the size of an effect. Hence this approach gives disproportionate weight to large studies with small effects since large samples result in a larger number of significant findings than do small samples. In addition, Hedges and Olkin (1980) have shown that in vote-counting there is a high probability of failing to conclude that a positive effect exists when in fact it does. These authors have also concluded that the probability of making this Type II error may increase as the number of studies increases! Thus in spite of its quantitative nature, the vote-counting approach can lead to misleading results and is generally not recommended for research reviews.

Combining Significance Levels

The classic chapter by Mosteller and Bush (1954) demonstrates that quantitative techniques for research integration have been available for decades. In their chapter, which appeared in the 1954 edition of the Handbook of Social Psychology, Mosteller and Bush presented several methods for combining tests of significance, drawing on work by Fisher (1938); Stouffer, Suchman, DeVinney, Star, and Williams (1949); and Gordon, Loveland, and Cureton (1952). Mosteller and Bush also discussed the problem of nonindependence of studies, an issue which concerns reviewers using quantitative approaches today.

Some years later, Rosenthal (1978) compared a number of methods for combining the probabilities obtained in two or more studies and suggested when the various procedures should and should not be used. However, a combined probability estimate does not indicate the size of the effect. Thus Rosenthal (1978, 1984) has recommended including an effect size for each combined probability estimate (e.g., Glass's delta, Cohen's d, or a correlation coefficient). Rosenthal has also suggested that a confidence interval be reported for the estimated effect size.

Meta-Analytic Approaches

As defined by Glass (1976) in his classic article, meta-analysis is the statistical analysis of the analytic results from a number of independent studies. Perhaps the most frequently used meta-analytic approach is the one developed by Glass and his colleagues (Glass, McGaw, & Smith, 1981; Smith et al., 1980). This approach has been extensively used to integrate research findings in psychology and education. The unit of analysis in Glass's approach is a standardized mean difference called the effect size (or "delta"). This effect size is defined as the difference between the means of the experimental and the control groups divided by the standard deviation of the control group ($ES = [M_e - M_c]/SD_c$)[2]. Glass's procedure is convenient for studies which report means and standard deviations.

Another meta-analytic approach is that of Hunter, Schmidt, and Jackson (1982). This approach was derived in large part from work by Schmidt, Hunter, and others on the generalizability of personnel selection validity studies. Schmidt and his colleagues have demonstrated that much of the variance in results across validity studies is due to statistical artifacts rather than to characteristics unique to the various settings as previously assumed. When corrections are made for artifacts such as sampling error, measurement error, and range restriction, the variation among studies shrinks dramatically. Accordingly, Hunter et al. (1982) have recommended that research integrators correct their effect sizes (correlation coefficients, d, or delta) and the associated variances for statistical artifacts (especially sampling error) before beginning to search for moderator variables that may be responsible for differences obtained across studies. The Hunter et al. (1982) approach has been applied primarily to correlational studies.

ADVANTAGES OF META-ANALYTIC APPROACHES

When carefully adhered to, quantitative approaches (especially the meta-analytic) yield results which are far more objective and replicable than those of traditional procedures for research integration. Cooper (1982, 1984) has emphasized that research review methodology should be just as rigorous as that employed by primary researchers. Jackson (1980) has criticized the failure of social scientists to develop an explicit

---

[2]Glass's "delta" is similar to Cohen's (1977) d. The Cohen statistic uses the pooled or within-group standard deviation as the divisor instead of the control group standard deviation.

methodology for research integration, with the outcome that "each review is the result of implicit methods, consciously or unconsciously selected by the reviewer" (p. 440).

While more rigorous approaches do not eliminate subjective judgment calls (which studies to include, which study characteristics to code, and so on), they do require the reviewer to be explicit in stating what decision rules were used in making such choices. It is then possible for other researchers to critique both the conclusions and the process the reviewer used in reaching those conclusions. Fiske (1983) has asserted that the explicitness of meta-analytic work "generates pressures to conceptualize and theorize more sharply" (p. 66). That is, the researcher must determine more precisely what the goals are, how they can be assessed, what the critical characteristics of the intervention are, etc. Although Fiske (1983) made specific reference to psychotherapy, his comments apply to other topical areas of interest to military psychologists.

One of the factors contributing to objectivity is the use of statistical procedures. The results of studies to be integrated are converted to a common metric. Each resulting effect size is then a datum suitable for statistical analysis. The use of statistical procedures also makes it possible to detect modest results that otherwise might be overlooked and to investigate effects which cannot be examined in a single study (e.g., effect of year of publication or gender of author). The findings of statistical analyses are more readily confirmed (or disconfirmed) than are subjective conclusions reached through human information processing.

Without the application of statistical procedures, the cognitive overload is formidable for all but a small number of studies. Indeed, Cooper and Rosenthal (1980) have demonstrated the superiority of a statistical approach for a set of only seven studies. For the plethora of studies available for many military psychological areas, quantitative research integration would seem essential.

Initial applications of meta-analytic procedures emphasized the derivation of general summary statements, which typically involved main effects. Cotton and Cook (1982) were among those who criticized earlier meta-analyses for ignoring interactions. But interactions can be coded and their effects examined. However, as Green and Hall (1984) have pointed out, there may not be enough studies investigating the same or similar interactions to carry out such a test.

Proper coding also permits examining the relationship of methodological variables to study outcomes. Rather than excluding studies which do not attain some subjective level of methodological rigor, the reviewer can code degree of randomization, reactivity of measures, etc. and then investigate their relationship to effects. In this way, exclusion of studies can be made on an empirical rather than a subjective basis.

An additional benefit of effect sizes (Glass's delta, Cohen's d, or correlation coefficients) is that they take magnitude of effect into account. The narrative, vote-counting, and combining probabilities approaches lack this advantage.

In sum, quantitative approaches are more systematic, more objective, and more replicable than traditional approaches to research integration. For a sizable body of literature, they provide the only practical means for integrating research results in a meaningful way. The meta-analytic approaches have the additional advantage of providing a measure of the magnitude of the effect.

PROBLEMS POSED BY META-ANALYTIC APPROACHES

While meta-analytic procedures constitute a powerful tool for integrating a body of research, a commitment to such procedures should not be undertaken lightly. Military psychologists who are contemplating the use of a meta-analytic approach should be aware of some of the problems involved. One encounters some of these problems with any method of research integration. Other problems are unique to, or at least exacerbated by, the use of a quantitative approach.

Identifying Pertinent Studies

One of the contributions to research integration made by Glass and his colleagues (e.g., Glass et al., 1981; Smith et al., 1980) has been their emphasis on collecting all studies bearing on a given topic. Often research reviewers have included only published studies or have excluded relevant studies they believed lacked methodological rigor. Since all studies are flawed methodologically in some way, Glass has argued that one should code methodological characteristics and investigate whether or not methodological imperfections are related to the results.

The reviewer may find it difficult, however, to obtain unpublished studies. Most American and Canadian dissertations are available through University Microfilms in Ann Arbor, but expense may be a factor here. Government reports can sometimes be located in data bases such as the National Technical Information Service (NTIS) and Defense Technical Information Center (DTIC). Conference papers present a real dilemma unless one has followed the topical area closely for some time and has a network of contacts in the field who exchange papers. Few organizations now seem to publish proceedings of their meetings, so this source of pertinent literature is limited.

The difficulty of locating unpublished studies leads to a consideration of what Rosenthal (1979, 1980) calls the "file drawer problem." It is possible that our journals publish a biased sample of the research which is actually conducted. Carrying this view to an extreme, one could argue that journals publish the 5% of the population of studies that contain significant results, while the 95% of the studies which resulted in nonsignificant findings repose in researchers' file drawers.

Rosenthal (1979, 1980) has presented a "fail-safe N" formula that can be used to estimate the number of studies averaging null results that would be required to negate the reviewer's findings. If a relatively few unfound studies would accomplish this result, then the reviewer must be wary about his or her conclusions. If, on the other hand, a great many studies would be required to nullify the findings, the reviewer can be more confident of the results obtained. Rosenthal's "fail-safe N" formula is applicable to correlational data. Orwin (1983) has developed an analogous formula appropriate for effect sizes such as Glass's delta or Cohen's d.

Selecting an Approach

In selecting an approach to research integration, the reviewer should be aware that different approaches are not necessarily mutually exclusive. One could, for example, supplement a meta-analytic treatment of studies with a narrative review of selected studies. In fact, Light and Pillemer (1984) have suggested blending quantitative and qualitative treatments of review data. Some studies may not contain sufficient data to be quantified, but they may offer useful information or insights that the reviewer wishes to include. Each review situation differs, and the reviewer must use his or her judgment in making such decisions.

The number of studies available for review and their homogeneity may indicate which integrative approach to use. Cook and Leviton (1980) compared traditional and meta-analytic reviews of essentially the same literature (self-serving attributions in studies of interpersonal influence). These authors concluded that meta-analysis is the preferred approach when the number of studies is large but is inappropriate for a small number of studies which are "heterogeneous with respect to methods and constructs" (p. 461).

The type of data contained in the research reports may suggest that one approach is more practical than another. If analysis of variance is used, as is the case with many experimental studies, then calculating effect sizes may be in order because usually (though not always) means and standard deviations are reported. Man-machine interface research (for example, on computer displays) typically involves two or more groups which are compared in terms of mean error rates, time, etc. If correlational data predominate in the research to be integrated, as is the case in military selection research, the researcher may wish to consider the correlational approach suggested by Hunter et al. (1982). It is also possible to convert effect sizes to correlation coefficients and vice versa (Hunter et al., 1982, p. 98).

Coding Studies

Coding studies is usually an arduous task, but the process will proceed more easily if the research review is planned in a rigorous fashion. Selecting appropriate variables and their levels to be coded requires a great deal of thought. Previous research may suggest pertinent variables, but explicit definitions of these variables and their levels may need to be developed. The topical area must be carefully defined--what is to be included and excluded, as well as the logic upon which the decisions are based, must be explicated. If the topic is "senior level leadership," for

example, what is "senior level"? And what is a leader? Is the research on civilians to be included, or is the search to be limited to military leadership? Coding conventions will provide guidance for coders, but coders need training and perhaps retraining. Jackson (1980), in his discussion of the problem of achieving reliable and valid coding of study characteristics, has recommended periodic assessment of intercoder reliability in cases in which coding is carried out over a lengthy period of time.

Sometimes variables that seem reasonable prove to be impractical for coding. The reliability of dependent measures is an interesting study characteristic that could be related to study outcomes. If, however, reliability data are rarely reported for the instruments used, no conclusions can be drawn. (Although such a reporting deficiency is, in and of itself, a finding that has implications for future research.)

Failure to report reliability data for measures complicates correcting for measurement error, as recommended by Hunter et al. (1982). If these data are not reported, one must estimate the reliabilities or undertake strenuous efforts to determine what they are. Orwin and Cordray (1985) reported they attempted to obtain reliabilities of outcome measures from the Mental Measurements Yearbook (Buros, 1938-1978) with "generally unsuccessful results" (p. 144). The author's experience in a similar search for reliability data was that the manuals for standardized instruments reported several reliability coefficients, and it was not always apparent which one would be the most appropriate.

Intercoder reliability also needs to be established. There are several ways to calculate intercoder reliability, and authorities do not agree on which is the best method. For example, Bullock and Svyantek (1985) recommended using percent of agreement for nominal variables when data distributions are skewed, while Krippendorff (1980) has held that the use of percent of agreement is "wholly deceptive" (p. 135) because it does not take chance agreement into account. Commonly used measures of reliability for continuous variables are the intraclass correlation coefficient (Winer, 1971) or Pearson's r. Cohen's kappa (Cohen, 1960; Reynolds, 1977) is frequently used for nominal variables as it indicates the degree of agreement beyond chance. See Krippendorff (1980) for other possibilities.

Jones, Johnson, Butler, and Main (1983) concluded that sometimes more than one reliability measure should be reported. These authors applied six frequently used indices of interrater agreement to the same data and reported substantial variations among the indices under certain conditions. Factors to be considered in selecting measures of intercoder agreement are type of data (continuous/nominal variables) and type of data distribution (high/low variation of ratings and high/low agreement among raters).

Analyzing Data

It has generally been held that each effect size constitutes a datum. Accordingly, the usual statistical analyses can be conducted just as in primary research (Glass et al., 1981). Hedges and Olkin (1985), however, have argued that applying conventional statistical procedures to effect size

data may involve violations of the assumptions underlying conventional analyses. Hedges and Olkin (1985) have asserted that conventional analyses are "problematic for both statistical and conceptual reasons" (p.9) and have presented arguments for this point of view. With respect to data analysis, the research integrator should also consider the aspects of nonindependence, outliers, and missing data.

Nonindependence. Over 30 years ago, Mosteller and Bush (1954) noted the lack of independence among studies whose results are to be quantitatively combined. For example, research efforts from the same laboratory or conducted by the same investigator cannot be considered truly independent of one another. With a meta-analytic approach, more than one effect size is typically calculated for each study.[3] For certain analyses, then, an individual study may contribute more than one effect size for the same group of subjects. Green and Hall (1984) have observed that reviewers are divided about whether they should use one result per group or whether it is permissible to use more than one. Each research integrator must consider the degree of nonindependence in the data and exercise his or her best judgment in dealing with the problem.

Outliers. When research results are quantified, some effect sizes (or correlations) will fall outside the usual range of outcomes. Barnett and Lewis (1978) have suggested that one may wish to examine outliers to identify unusual patterns or to delete extreme cases if one is primarily interested in overall results. Light and Pillemer (1984) have emphasized that it is in the tails of distributions that the reviewer may discover the outstandingly good or poor interventions or individuals that can clarify the research thrust under investigation. In their book on statistical methods for meta-analysis, Hedges and Olkin (1985) have described both graphic and statistical methods for identifying outliers. Tukey (1977) and Hartwig and Dearing (1979), in their book on exploratory data analysis, also have presented procedures for detecting outliers. Once outliers are identified, however, the research integrator must use his or her judgment in deciding what to do about them. One approach is to report the results with and without the outliers.

Missing data. The data reported by researchers in their articles and reports are frequently incomplete. Many authors have commented on the problem of inadequate reporting (e.g, Green & Hall, 1984; Jackson, 1980; Orwin & Cordray, 1985). Light and Pillemer (1984) noted that the inadequacy of research reporting is "surprising" and that some might wish to "change the word surprising to shocking" (p. 101). After experiencing the difficulties of integrating some of the career counseling outcome literature, Oliver and Spokane (1983) described the problem of inadequate research reporting and suggested guidelines for such reporting. More recently, the new editor of

---

[3]Using Glass's approach, an effect size is computed for each comparison of an experimental group with a control group on each dependent variable. Hence, for a study with two experimental groups and three dependent variables, six effect sizes would be calculated.

the Academy of Management Journal included in her guidelines for authors the requirement that data required for meta-analytic research reviews be included in all articles submitted to the journal (Beyer, 1985).

In quantifying research results, means and standard deviations (or correlations) are required for both significant and nonsignificant results. Since many research reports fail to report these basic data (as well as n's, exact probability levels, etc.), the reviewer must decide what to do in cases of missing data. Books such as Glass et al. (1981) and Smith et al. (1980) provide guidance in this area. Orwin and Cordray (1985) have discussed the effects of deficient reporting on meta-analytic results. Authorities may differ, however, in what to do in various cases of missing data. In such instances, the reviewer must make a judgment call. It is important to document one's rationale for such decisions and to apply the decision rule consistently.

Reporting Results

Quantitative versus qualitative results. Too often in primary research, it has been assumed that researchers had to choose between quantitative techniques or qualitative methods. More recently (Light & Pillemer, 1984; Van Maanen, Dabbs, & Faulkner, 1982), it has been emphasized that this is not an either-or situation. There is not only room but, indeed, a need for a variety of research approaches, perhaps even within one research project (Rossman & Wilson, 1985). And so it also is with research integration. Light and Pillemer (1984) argued cogently for (and presented pertinent examples of) the combination of what they call "numbers and narrative." Perhaps the reviewer has found articles or books which lack data (as, for instance, in the leadership literature), but believes that these writings contain useful insights or suggest research directions for the topical area. Rather than omit the contributions such writings could make to the research area being reviewed, the reviewer can incorporate their findings by means of a narrative exposition.

Displaying data. During the last few years, some interesting and innovative ways of displaying data have been demonstrated. Light and Pillemer (1984) have shown how a simple frequency distribution of effect sizes can reveal the orderliness of what might otherwise appear to be a chaotic and confusing set of research results. These authors also have suggested the use of "funnel displays" to identify publication bias and historical trends. Some of Tukey's (1977) techniques (e.g., the "jackknife" method and "stem and leaf" displays) can also be used in displaying data from quantitative reviews.

Another interesting method for displaying data is the Binomial Effect Size Display (BESD). Developed by Rosenthal and Rubin (1982), the BESD is a means of depicting the change in a success rate (such as an improvement rate, selection rate, or survival rate) which can be attributed to a new selection procedure, predictor variable, or intervention. An example from Rosenthal (1984, p. 131) will illustrate the use of this technique: For an estimated mean effect size corresponding to an r of .30, the increase in success rate would be from 35% to 65%. If, for example, this effect were associated with

a basic skills remedial program, it would mean that the pass rate of persons participating in the program would rise from 35% to 65%. By using the BESD, one can demonstrate that rather modest effects can have substantial importance in the real world.

The BESD is particularly appropriate for demonstrating to policy makers the potential effects of possible changes. Military psychologists frequently need to translate their research results into real-world effects, and techniques such as BESD represent a simple and effective means of doing so. When based on a number of studies investigating essentially the same problem, the BESD provides a very solid basis upon which to base policy changes (or to reject such changes), for example.

A final caveat. Cooper and Rosenthal (1980) found that subjects in their quantitative review condition reached stronger conclusions than those in the traditional review condition. As Green and Hall (1984) have pointed out, this outcome is to be expected. Quantitative reviews generally involve a greater number of studies, and statistical methods make it possible to detect effects that might otherwise be overlooked.

But the quantitative nature of meta-analysis makes it all the more necessary for research integrators to be meticulous in their use of this approach. In reporting the results of a meta-analysis, the reviewer must scrupulously describe the methodology. It is particularly important to explain what was done in cases of missing data. If too many effect sizes must be estimated (rather than calculated directly from means and standard deviations), one's certainty about the results would certainly be diminished. Use of procedures such as a fail-safe N and attention to factors such as nonindependence and outliers will enable the reviewer and the reader to attain a balanced perspective.

IMPLICATIONS FOR MILITARY PSYCHOLOGY

The increasing attention on research integration has important implications for military psychologists. Some of these are discussed below.

More Complete Reporting of Research Results

Anyone who has attempted to use a meta-analytic approach to research integration has invariably been appalled at the inadequacies of research reporting. Sometimes a quantitative approach must be abandoned. For example, a colleague who wished to apply a meta-analytic approach to the research on senior leadership found that only about 10 percent of the 63 studies which had been identified in this area involved comparison groups and contained data sufficient for the use of such an approach (A. G. Steinberg, personal communication, September 26, 1985).

In the opinion of the writer, there is no substitute for providing complete data in a research report, be it article or institutional report. Obtaining additional data from authors is generally unsatisfactory as they may move, die, or be unable to find the needed data. At a minimum, authors should report means and standard deviations (or correlation coefficients),

n's for each group, clear descriptions of samples and procedures, and intercorrelations of quantitative variables. Reliability and validity data on measures used should also be routinely reported. As has been documented (e.g., Oliver, 1979; Orwin & Cordray, 1985), reliability and validity data are infrequently reported, especially for author-constructed measures.

Military psychologists must be sure that their own research reports contain the data needed for quantitative research integration. In addition, when military psychologists review papers as journal editors, as members of editorial boards, or as technical reviewers, they should recommend for publication (in journals or institutional reports) only those manuscripts which meet the minimum data reporting standards. In institutional reports intended for a military audience, such details can be contained in a technical supplement. For journal articles or technical reports, additional data can be placed in an appendix. In any event, it is essential that complete data be reported.

Integration of Research in Appropriate Topical Areas

Military psychologists tend, as do their civilian counterparts, "to reinvent the wheel." That is, previously used instruments that measure the same variable or previous investigations of essentially the same problem are not considered in planning new research. It is very important that military psychologists integrate the research findings of the myriad of investigations which have been conducted over the years. We need to develop the "cumulative knowledge" that Schmidt (1980) found lacking. Take the training area, for instance. There must be thousands of training studies which could be quantitatively integrated--not in one gigantic meta-analysis, of course, but broken down into appropriate and manageable sub-areas. Some of the topical areas of interest to military psychologists have stimulated both civilian and military research (e.g., leadership). A comparison of the civilian and military research could be valuable.

In attempting to integrate large bodies of research of interest to the military, it would be beneficial to coordinate such projects across the services. Such an effort is now being considered. (E. Salas, personal communication, September 11, 1985). Although collaborative efforts take considerable time and effort to effect, the potential payoff is great.

Identification of Gaps in Military Psychological Research

In addition to finding out what we already do know, we need to find out what we do not know. That is, where are the gaps in our knowledge of a given area? Quantitative integration of available research can clarify the areas that need additional investigation. Perhaps we can make some generalizations about main effects in a certain research area, but it may be the interactions involved that will be most important to us. We might know, for example, that a certain type of training is effective. But for whom is this type of training most successful? Under what specific circumstances is it successful? Although research to date may indicate that the training is

generally effective, it may be the identification of the moderator variables involved that will enable our trainers to make the most effective use of training time.

Here, again, collaborative efforts by military psychologists across the services would be profitable for all concerned. Well planned programmatic research could lead to advances that would be far greater than the sum of our individual efforts.

REFERENCES

Barnett, V., & Lewis, T. (1978). Outliers in statistical data. Chichester, U.K.: Wiley.

Beyer, J. M. (1985). From the editor. Academy of Management Journal, 28, 5-8.

Bullock, R. J., & Svyantek, D. J. (1985). Analyzing meta-analysis: Potential problems, an unsuccessful replication, and evaluation criteria. Journal of Applied Psychology, 70, 108-115.

Buros, O. K. (1938-1978). Mental measurements yearbook. Highland Park, NJ: Gryphon Press.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. New York: Academic Press.

Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. Journal of Personality, 48, 449-472.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. Review of Educational Research, 5, 3-8.

Cooper, H. M. (1984). The integrative research review: A systematic approach. Beverly Hills, CA: Sage

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. Psychological Bulletin, 87, 442-449.

Cotton, J. L., & Cook, M. S. (1982). Meta-analyses and the effects of various reward systems: Some different conclusions from Johnson et al. Psychological Bulletin, 92, 176-183.

Fisher, R. A. (1938). Statistical methods for research workers (7th ed.). London: Oliver & Boyd.

Fiske, D. W. (1983). The meta-analytic revolution in outcome research. Journal of Consulting and Clinical Psychology, 51, 65-70.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G. B. (1977). Integrating findings: The meta-analysis of research. Review of Research in Education, 5, 351-379.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Gordon, M. H., Loveland, L. H., & Cureton, E. E. (1952). An extended table of chi-square for two degrees of freedom, for use in combining probabilities from independent samples. Psychometrika, 17, 311-316.

Green, B. F., & Hall, J. A. (1984). Cumulative methods for literature reviews. Annual Review of Psychology, 35, 37-53.

Hartwig, F., & Dearing, B. E. (1979). Exploratory data analysis. Beverly Hills, CA: Sage.

Hedges, L. V., & Olkin, I. (1980). Vote-counting methods in research synthesis. Psychological Bulletin, 88, 359-369.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). Meta-analysis: Cumulating research findings across studies. Beverly Hills, CA: Sage.

Jackson, G. B. (1980). Methods for integrative reviews. Review of Educational Research, 50, 438-460.

Jones, A. P., Johnson, L. A., Butler, M. C., & Main, D. S. (1983). Apples and oranges: An empirical comparison of commonly used indices of interrater agreement. Academy of Management Journal, 26, 507-519.

Krippendorff, K. (1980). Content analysis: An introduction to its methodology. Beverly Hills, CA: Sage.

Light, R. J., & Pillemer, D. B. (1984). Summing up: The science of reviewing research. Cambridge, MA: Harvard University Press.

Mosteller, F., & Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), Handbook of social psychology: Vol. I, Theory and method. (pp. 289-334). Cambridge, MA: Addison-Wesley.

Oliver, L. W. (1979). Outcome measurement in career counseling research. Journal of Counseling Psychology, 26, 217-226.

Oliver, L. W. (1984). Military psychology (pp. 375-376). In R. Corsini (Ed.), Encyclopedia of psychology. New York: Wiley.

Oliver, L. W., & Spokane, A. R. (1983). Research integration: Approaches, problems, and recommendations for research reporting. Journal of Counseling Psychology, 30, 252-257.

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. Journal of Educational Statistics, 8, 157-159.

Orwin, R. G., & Cordray, D. S. (1985). Effects of deficient reporting on meta-analysis: A conceptual framework and reanalysis. Psychological Bulletin, 97, 134-147.

Peters, L. H., Hartke, D. D., & Pohlman, J. T. (1985). Fiedler's contingency theory of leadership: An application of the meta-analysis procedures of Schmidt and Hunter. Psychological Bulletin, 97, 274-285.

Reynolds, H. T. (1977). The analysis of cross-classifications. New York: The Free Press.

Rosenthal, R. (1978). Combining results of independent studies. Psychological Bulletin, 85, 185-193

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. Psychological Bulletin, 86, 638-641.

Rosenthal, R. (1980). Summarizing significance levels. In R. Rosenthal (Ed.), New directions for methodology of social and behavioral science: No. 5. Quantitative assessment of research domains (pp. 33-46). San Francisco: Jossey-Bass.

Rosenthal, R. (1984). Meta-analytic procedures for social research. Beverly Hills, CA: Sage.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. Journal of Educational Psychology, 74, 166-169.

Rossman, G. G., & Wilson, B. L. (1985). Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. Evaluation Review, 9, 627-643.

Schmidt, F. L. (1980, September). The research task of the 1980s: Integrating research findings across studies to produce cumulative knowledge. Paper presented at the meeting of the American Psychological Association, Montreal.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). The benefits of psychotherapy. Baltimore, MD: John Hopkins University Press.

Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S. A., & Williams, R. M. Jr. (1949). The American soldier: Vol. 1. Adjustment during Army life.

Strube, M. J., & Garcia, J. E. (1981). A meta-analytic investigation of Fiedler's contingency model of leadership effectiveness. Psychological Bulletin, 90, 307-321.

Tukey, J. W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.

Van Maanen, J., Dabbs, J. M. Jr., & Faulkner, R. R. (1982). Varieties of qualitative research. Beverly Hills, CA: Sage.

Winer, B. J. (1971). Statistical principles in experimental design. New York: McGraw-Hill.